



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : H04R 25/00, A61F 2/18	A1	(11) International Publication Number: WO 91/03913
		(43) International Publication Date: 21 March 1991 (21.03.91)

(21) International Application Number: PCT/AU90/00407

(22) International Filing Date: 7 September 1990 (07.09.90)

(30) Priority data:
PJ 6249 8 September 1989 (08.09.89) AU

(71) Applicants: COCHLEAR PTY. LIMITED [AU/AU]; 14 Mars Road, Lane Cove, NSW 2066 (AU). THE UNIVERSITY OF MELBOURNE [AU/AU]; Grattan Street, Parkville, VIC 3052 (AU).

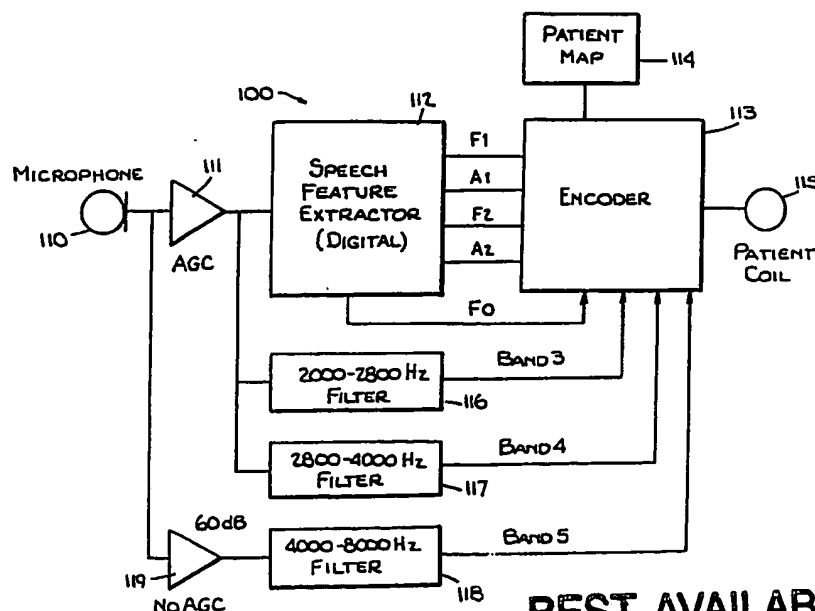
(72) Inventors: SELIGMAN, Peter, Misha ; 107 Fawkner Street, Essendon, VIC 3040 (AU). DOWELL, Richard, C. ; 169 Progress Road, North Eltham, VIC 3095 (AU).

(74) Agent: MAXWELL, Peter, Francis; Peter Maxwell & Associates, Blaxland House, 5 Ross Street, North Parramatta, NSW 2151 (AU).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent)*, DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
With amended claims.

(54) Title: MULTI-PEAK SPEECH PROCESSOR



(57) Abstract

BEST AVAILABLE COPY

An improved pulsatile system for a cochlear prosthesis is disclosed. The system employs a multi-spectral peak coding strategy to extract a number, for example five, of spectral peaks from an incoming acoustic signal received by a microphone. It encodes this information into sequential pulses that are sent to selected electrodes of a cochlear implant. The first formant (F1) spectral peak (280-1000 Hz) and the second formant (F2) spectral peak (800-4000 Hz) are encoded and presented to apical and basal electrodes, respectively. F1 and F2 electrode selection follows the tonotopic organization of the cochlea. High-frequency spectral information is sent to more basal electrodes and low-frequency spectral information is sent to more apical electrodes. Spectral energy in the regions of 2000-2800 Hz, 2800-4000 Hz, and above 4000 Hz is encoded and presented to three fixed electrodes. The fundamental or voicing frequency (F0) determines the pulse rate of the stimulation during voiced periods and a pseudo-random aperiodic rate determines the pulse rate of stimulation during unvoiced periods. The amplitude of the acoustic signal in the five bands determines the stimulus intensity.

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Fasso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	PL	Poland
CA	Canada	JP	Japan	RO	Romania
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	LJ	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
DE	Germany	LU	Luxembourg	TD	Chad
DK	Denmark			TG	Togo
				US	United States of America

MULTI-PEAK SPEECH PROCESSOR

Technical Field

This invention relates to pulsatile type multi-channel cochlear implant systems for the totally or profoundly deaf.

Background of the Invention

Pulsatile multi-channel cochlear implant systems generally include a cochlear implant, an external speech processor, and an external headset. The cochlear implant delivers electrical stimulation pulses to an electrode array (e.g., 22 electrodes) placed in the cochlea. The speech processor and headset transmit information and power to the cochlear implant.

The speech processor operates by receiving an incoming acoustic signal from a microphone in the headset, or from an alternative source, and extracting from this signal specific acoustic parameters. Those acoustic parameters are used to determine electrical stimulation parameters, which are encoded and transmitted to the cochlear implant via a transmitting coil in the headset, and a receiving coil forming part of the implant.

In many people who are profoundly deaf, the reason for deafness is absence of, or destruction of, the hair cells in the cochlea which transduce acoustic signals into nerve impulses. These people are thus unable to derive any benefit from conventional hearing aid systems, no matter how loud the acoustic stimulus is made, because it is not possible for nerve impulses to be generated from sound in the normal manner. Cochlear implant systems seeks to bypass these hair cells by presenting electrical stimulation to the auditory nerve fibers directly, leading to the perception of sound in the brain. There have been many ways

described in the past for achieving this object, running from implantation of electrodes in the cochlea connected to the outside world via a cable and connector attached to the patient's skull, to sophisticated multichannel devices communicating with an external computer via radio frequency power and data links.

The invention described herein is particularly suited for use in a prosthesis which comprises a multichannel electrode implanted into the cochlea, connected to a multichannel implanted stimulator unit, which receives power and data from an externally powered wearable speech processor, wherein the speech processing strategy is based on known psychophysical phenomenon, and is customized to each individual patient by use of a diagnostic and programming unit. One example of such a prosthesis is the one shown and described in U.S. Patent No. 4,532,930 to Crosby et al., entitled "Cochlear Implant System for an Auditory Prosthesis."

In order to best understand the invention it is necessary to be aware of some of the physiology and anatomy of human hearing, and to have a knowledge of the characteristics of the speech signal. In addition, since the hearing sensations elicited by electrical stimulation are different from those produced by acoustic stimulation in a normal hearing person, it is necessary to discuss the psychophysics of electrical stimulation of the auditory system. In a normal hearing person, sound impinges on the ear drum, as illustrated in FIG. 1, and is transmitted via a system of bones called the ossicles, which act as levers to provide amplification and acoustic impedance matching to a piston, or membrane, called the oval window, which is coupled to the cochlea chamber.

The cochlear chamber is about 35 mm long when unrolled and is divided along most of its length by a partition. This partition is called the basilar membrane. The lower chamber is called the scala tympani. An opening at the remote end of the cochlea chamber communicates between the upper and lower halves thereof. The cochlea is filled with a fluid having a viscosity of about twice that of water. The scala tympani is provided with another piston or membrane called the round window which serves to take up the displacement of the fluid.

When the oval window is acoustically driven via the ossicles, the basilar membrane is displaced by the movement of fluid in the cochlea. By the nature of its mechanical properties, the basilar membrane vibrates maximally at the remote end or apex of the cochlea for low frequencies, and near the base or oval window thereof for high frequencies. The displacement of the basilar membrane stimulates a collection of cells called the hair cells situated in a special structure on the basilar membrane. Movements of these hairs produce electrical discharges in fibers of the VIIIth nerve, or auditory nerve. Thus the nerve fibers from hair cells closest to the round window (the basal end of the cochlea) convey information about high frequency sound, and fibers more apical convey information about low frequency sound. This is referred to as the tonotopic organization of nerve fibers in the cochlea.

Hearing loss may be due to many causes, and is generally of two types. Conductive hearing loss occurs when the normal mechanical pathways for sound to reach the hair cells in the cochlea are impeded, for example by damage to the ossicles. Conductive hearing loss may often be helped by use of hearing

aids, which amplify sound so that acoustic information does reach the cochlea. Some types of conductive hearing loss are also amenable to alleviation by surgical procedures.

Sensorineural hearing loss results from damage to the hair cells or nerve fibers in the cochlea. For this type of patient, conventional hearing aids will offer no improvement because the mechanisms for transducing sound energy into nerve impulses have been damaged. It is by directly stimulating the auditory nerve that this loss of function can be partially restored.

In the system described herein, and in some other cochlear implant systems in the prior art, the stimulating electrodes are surgically placed in the scala tympani, in close proximity to the basilar membrane, and currents that are passed between the electrodes result in neural stimulation in groups of nerve fibers.

The human speech production system consists of a number of resonant cavities, the oral and the nasal cavities, which may be excited by air passing through the glottis or vocal cords, causing them to vibrate. The rate of vibration is heard as the pitch of the speaker's voice and varies between about 100 and 400 Hz. The pitch of female speakers is generally higher than that of male speakers.

It is the pitch of the human voice which gives a sentence intonation, enabling the listener, for instance, to be able to distinguish between a statement and a question, segregate the sentences in continuous discourse and detect which parts are particularly stressed. This together with the amplitude of the signal provides the so-called prosodic information.

Speech is produced by the speaker exciting the vocal cords, and manipulating the acoustic cavities by movement of the tongue, lips and jaw to produce different sounds. Some sounds are produced with the vocal cords excited, and these are called voiced sounds. Other sounds are produced by other means, such as the passage of air between teeth and tongue, to produce unvoiced sounds. Thus the sound "Z" is a voiced sound, whereas "S" is an unvoiced sound; "B" is a voiced sound, and "P" is an unvoiced sound, etc.

The speech signal can be analyzed in several ways. One useful analysis technique is spectral analysis, whereby the speech signal is analyzed in the frequency domain, and a spectrum is considered of amplitude (and phase) versus frequency. When the cavities of the speech production system are excited, a number of spectral peaks are produced, and the frequencies and relative amplitudes of these spectral peaks are also varied with time.

The number of spectral peaks ranges between about three and five and these peaks are called "formants". These formants are numbered from the lowest frequency formant, conventionally called F1, to the highest frequency formants, and the voice pitch is conventionally referred to as F0. Characteristic sounds of different vowels are produced by the speaker changing the shape of the oral and nasal cavities, which has the effect of changing the frequencies and relative intensities of these formants.

In particular, it has been found that the second formant (F2) is important for conveying vowel information. For example, the vowel sounds "oo" and "ee" may be produced with identical voicings of the vocal cords, but will sound different due to different second formant characteristics.

There is of course a variety of different sounds in speech and their method of production is complex. For the purpose of understanding the invention herein however, it is sufficient to remember that there are two main types of sounds - voiced and unvoiced; and that the time course of the frequencies and amplitudes of the formants carries most of the intelligibility of the speech signal.

The term "psychophysics" is used herein to refer to the study of the perceptions elicited in patient's by electrical stimulation of the auditory nerve. For stimulation at rates between 100 and 400 pulses per second; a noise is perceived which changes pitch with stimulation rate. This is such a distinct sensation that it is possible to convey a melody to a patient by its variation.

By stimulating the electrode at a rate proportional to voice pitch (F_0), it is possible to convey prosodic information to the patient. This idea is used by some cochlear implant systems as the sole method of information transmission, and may be performed with a single electrode.

It is more important to convey formant information to the patient, as this contains most of the intelligibility of the speech signal. It has been discovered by psychophysical testing that just as an auditory signal which stimulates the remote end of the cochlea produces a low frequency sensation and a signal which stimulates the near end thereof produces a high frequency sensation, a similar phenomenon will be observed with electrical stimulation. The perceptions elicited by electrical stimulation at different positions inside the cochlea have been reported by the subjects as producing percepts which vary in "sharpness" or "dullness", rather than pitch as such. How-

ever, the difference in frequency perceptions between electrodes is such that formant, or spectral peak, information can be coded by selection of electrode, or site of stimulation in the cochlea.

It has been found by psychophysical testing that the perceived loudness of sounds elicited by electrical stimulation of the auditory nerve has a larger dynamic range than the dynamic range of the stimulation itself. For example, a 220 dB dynamic range of electrical stimulation may produce perceptions from threshold or barely perceivable, to threshold of pain. In normal hearing people the dynamic range of sound perception is in the order of 100 dB.

It has also been discovered through psychophysical testing that the pitch of sound perceptions due to electrical stimulation is also dependent upon frequency of stimulation, but the perceived pitch is not the same as the stimulation frequency. In particular, the highest pitch able to be perceived through the mechanism of the changing stimulation rate alone is in the order of 1 kHz, and stimulation at rates above this maximum level will not produce any increase in frequency or pitch of the perceived sound. In addition, for electrical stimulation within the cochlea, the perceived pitch depends upon electrode position. In multiple electrode systems, the perceptions due to stimulation at one electrode are not independent of the perceptions due to simultaneous stimulation of nearby electrodes. Also, the perceptual qualities of pitch, "sharpness", and loudness are not independently variable with stimulation rate, electrode position, and stimulation amplitude.

Some systems of cochlear implants in the prior art are arranged to stimulate a number of electrodes simultaneously in proportion to the energy in

specific frequency bands, but this is done without reference to the perceptions due to stimulus current in nearby stimulating electrodes. The result is that there is interaction between the channels and the loudness is affected by this.

A number of attempts have heretofore been made to provide useful hearing through electrical stimulation of auditory nerve fibers, using electrodes inside or adjacent to some part of the cochlear structure. Systems using a single pair of electrodes are shown in U.S. Patent No. 3,751,605 to Michelson and U.S. Patent No. 3,752,939 to Bartz.

In each of these systems an external speech processing unit converts the acoustic input into a signal suitable for transmission through the skin to an implanted receiver/stimulator unit. These devices apply a continuously varying stimulus to the pair of electrodes, stimulating at least part of the population of auditory nerve fibers, and thus producing a hearing sensation.

The stimulus signal generating from a given acoustic input is different for each of these systems, and while some degree of effectiveness has been demonstrated for each, performance has varied widely across systems and also for each system between patients. Because the design of these systems has evolved empirically, and has not been based on detailed psychophysical observations, it has not been possible to determine the cause of this variability. Consequently, it has not been possible to reduce it.

An alternative approach has been to utilize the tonotopic organization of the cochlea to stimulate groups of nerve fibers, depending on the frequency spectrum of the acoustic signal. Systems using this technique are shown in U.S. Patent No. 4,207,441 to

Ricard, U.S. Patent No. 3,449,753 to Doyle, U.S. Patent No. 4,063,048 to Kissiah, and U.S. Patents No. 4,284,856 and No. 4,357,497 to Hochmair et al.

The system described by Kissiah uses a set of analog filters to separate the acoustic signal into a number of frequency components, each having a predetermined frequency range within the audio spectrum. These analog signals are converted into digital pulse signals having a pulse rate equal the frequency of the analog signal they represent, and the digital signals are used to stimulate the portion of the auditory nerve normally carrying the information in the same frequency range. Stimulation is accomplished by placing an array of spaced electrodes inside the cochlea.

The Kissiah system utilizes electrical stimulation at rates up to the limit of normal acoustic frequency range, say 10 kHz, and independent operation of each electrode. Since the maximum rate of firing of any nerve fiber is limited by physiological mechanisms to one or two kHz, and there is little perceptual difference for electrical pulse rates above 800 Hz, it may be inappropriate to stimulate at the rates suggested. No consideration is given to the interaction between the stimulus currents generated by different electrodes, which experience shows may cause considerable uncontrolled loudness variations, depending on the relative timing of stimulus presentations. Also, this system incorporates a percutaneous connector which has with it the associated risk of infection.

The system proposed by Doyle limits the stimulation rate for any group of fibers to a rate which would allow any fiber to respond to sequential stimuli. It utilizes a plurality of transmission channels, with each channel sending a simple composite

power/data signal to a bipolar pair of electrodes. Voltage source stimulation is used in a time multiplexed fashion similar to that subsequently used by Ricard and described below, and similar uncontrolled loudness variations will occur with the suggested independent stimulation of neighboring pairs of electrodes. Further, the requirement of a number of transmission links equal to the number of electrode pairs prohibits the use of this type of system for more than a few electrodes.

The system proposed by Ricard utilizes a filter bank to analyze the acoustic signal, and a single radio link to transfer both power and data to the implanted receiver/stimulator, which presents a time-multiplexed output to sets of electrodes implanted in the cochlea. Monophasic voltage stimuli are used, with one electrode at a time being connected to a voltage source while the rest are connected to a common ground line. An attempt is made to isolate stimulus currents from one another by placing small pieces of silastic inside the scala, between electrodes. Since monophasic voltage stimuli are used, and the electrodes are returned to the common reference level after presentation of each stimulus, the capacitive nature of the electrode/electrolyte interface will cause some current to flow for a few hundred microseconds after the driving voltage has been returned to zero. This will reduce the net transfer of charge (and thus electrode corrosion) but this charge recovery phase is now temporarily overlapped with the following stimulus or stimuli. Any spatial overlap of these stimuli would then cause uncontrolled loudness variations.

In the Hochmair et al. patents a plurality of carrier signals are modulated by pulses correspond-

ing to signals in audio frequency bands. The carrier signals are transmitted to a receiver having independent channels for receiving and demodulating the transmitted signals. The detected pulses are applied to electrodes on a cochlear implant, with the electrodes selectively positioned in the cochlea to stimulate regions having a desired frequency response. The pulses have a frequency which corresponds to the frequency of signals in an audio band and a pulse width which corresponds to the amplitude of signals in the audio band.

U.S. Patent No. 4,267,410 to Forster et al. describes a system which utilizes biphasic current stimuli of predetermined duration, providing a good temporal control of both stimulating and recovery phases. However, the use of fixed pulse duration prohibits variation of this parameter which may be required by physiological variations between patients. Further, the data transmission system described in this system severely limits the number of pulse rates available for constant rate stimulation.

U.S. Patent No. 4,593,696 to Hochmair et al. describes a system in which at least one analog electrical signal is applied to implanted electrodes in a patient, and at least one pulsatile signal is applied to implanted electrodes. The analog signal represents a speech signal, and the pulsatile signal provides specific speech features such as formant frequency and pitch frequency.

U.S. Patent No. 4,515,158 to Patrick et al. describes a system in which sets of electrical currents are applied to selected electrodes in an implanted electrode array. An incoming speech signal is processed to generate an electrical input corresponding to the received speech signal, and electrical

signals characterizing accoustic features of the speech signal are generated from the input signal. Programmable means obtains and stores data from the electrical signals and establishes sets of electric stimuli to be applied to the electrode array, and instruction signals are produced for controlling the sequential application of pulse stimuli to the electrodes at a rate derived from the voicing frequency of the speech signal for voiced utterances and at a rate independent of the voicing frequency for unvoiced utterances.

The state of the art over which the present invention represents an improvement is perhaps best exemplified by the aforesaid U.S. Patent No. 4,532,930 to Crosby et al., entitled "Cochlear Implant System for an Auditory Prosthesis". The subject matter of said Crosby et al. patent is hereby incorporated herein by reference. The Crosby et al. patent describes a cochlear implant system in which an electrode array comprising multiple platinum ring electrodes in a silastic carrier is implanted in the cochlea of the ear. The electrode array is connected to a multi-channel receiver-stimulating unit, containing a semiconductor integrated circuit and other components, which is implanted in the patient adjacent the ear. The receiver-stimulator unit receives data information and power through a tuned coil via an inductive link with a patient-wearable external speech processor. The speech processor includes an integrated circuit and various components which are configured or mapped to emit data signals from an Erasable Programmable Read Only Memory (EPROM). The EPROM is programmed to suit each patient's electrical stimulation perceptions, which are determined through testing of the patient and his implanted stimulator/electrode. The

testing is performed using a diagnostic and programming unit (DPU) that is connected to the speech processor by an interface unit.

The Crosby et al. system allows use of various speech processing strategies, including dominant spectral peak and amplitude compression of voice pitch, so as to include voiced sounds, unvoiced glottal sounds and prosodic information. The speech processing strategy employed is based on known psychophysical phenomenon, and is customized to each individual patient by the use of the diagnostic and programming unit. Biphasic pulses are supplied to various combinations of the electrodes by a switch controlled current sink in various modes of operation. Transmission of data is by a series of discrete data bursts which represent the chosen electrode(s), the electrode mode configuration, the stimulating current, and amplitude determined by the duration of the amplitude burst.

Each patient will have different perceptions resulting from electrical stimulation of the cochlea. In particular, the strength of stimulation required to elicit auditory perceptions of the same loudness may be different from patient to patient, and from electrode to electrode for the same patient. Patients also may differ in their abilities to perceive pitch changes from electrode to electrode.

The speech processor accommodates differences in psychophysical perceptions between patients and compensates for the differences between electrodes in the same patient. Taking into account each individual's psychophysical responses, the speech processor encodes acoustic information with respect to stimulation levels, electrode frequency boundaries, and other parameters that will evoke appropriate auditory

perceptions. The psychophysical information used to determine such stimulation parameters from acoustic signals is referred to as a MAP and is stored in a random access memory (RAM) inside the speech processor. An audiologist generates and "fine tunes" each patient's MAP using a diagnostic and programming system (DPS). The DPS is used to administer appropriate tests, present controlled stimuli, and confirm and record test results.

The multi-electrode cochlear prosthesis has been used successfully by profoundly deaf patients for a number of years and is a part of everyday life for many people in various countries around the world. The implanted part of the prosthesis has remained relatively unchanged except for design changes, such as those made to reduce the overall thickness of the device and to incorporate an implanted magnet to eliminate the need for wire headsets.

The external speech processor has undergone significant changes since early versions of the prosthesis. The speech coding scheme used by early patients presented three acoustic features of speech to implant users. These were amplitude, presented as current level of electrical stimulation; fundamental frequency or voice pitch, presented as rate of pulsatile stimulation; and the second formant frequency, represented by the position of the stimulating electrode pair. This coding scheme (F0F2) provided enough information for profoundly postlinguistically deafened adults to show substantial improvements in their perception of speech.

The early coding scheme progressed naturally to a later coding scheme in which additional spectral information is presented. In this scheme a second stimulating electrode pair was added, representing the

first formant of speech. The new scheme (FOF1F2) showed improved performance for adult patients in all areas of speech perception.

Despite success of speech processors using the FOF1F2 scheme over the last few years, a number of problems have been identified. For example, patients who perform well in quiet conditions can have significant problems when there is a moderate level of background noise. Also, the FOF1F2 scheme codes frequencies up to about 3,500Hz; however, many phonemes and environmental sounds have a high proportion of their energy above this range making them inaudible to the implant user in some cases.

According to one aspect of the invention there is provided an improved pulsatile system for a cochlear prosthesis in which an incoming audio signal is concurrently presented to a speech feature extractor and a plurality of band pass filters, the pass bands of which are different from one another and at least one of which is at a higher frequency than the normal range of the second formant or frequency peak of the speech signal. The energy within these pass bands controls the amplitude of electrical stimulation of a corresponding number of fixed electrode pairs adjacent the basal end of the electrode array, thus providing additional information about high frequency sounds at a tonotopically appropriate place within the cochlea. Preferably three additional band pass filters are employed in the ranges of 2,000 to 2,800Hz, 2,800 to 4,000Hz and 4,000 to

8,000Hz.

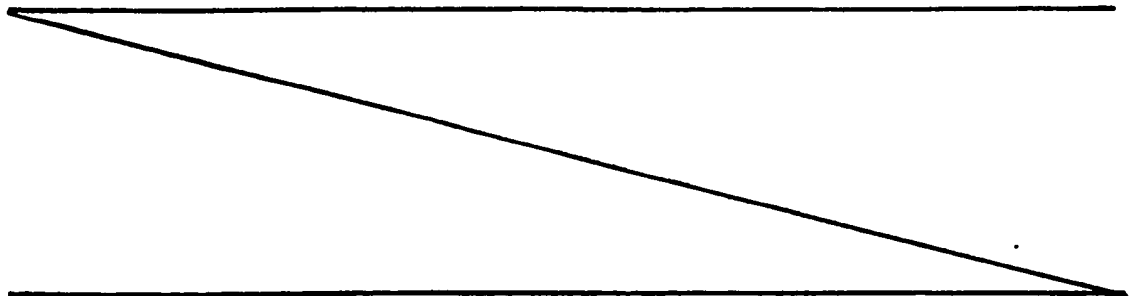
The overall stimulation rate remains as F0 (fundamental frequency or voice pitch) but, in addition, four electrical stimulation pulses occur for each glottal pulse, as compared with the F0F1F2 strategy heretofore used, in which only two pulses occur per voice pitch period. For voiced speech sounds, pulses representing the first and second formant are provided along with additional stimulation pulses representing energy in the 2,000 to 2,800Hz and the 2,800 to 4,000Hz ranges. For unvoiced phonemes, yet another pulse representing energy above 4,000Hz is provided while no stimulation for the first formant is provided, since there is no energy in this frequency range. Stimulation occurs at a random pulse rate of approximately 260Hz, which is about double that used in earlier speech coding schemes.

According to another aspect of the invention there is provided a method of processing an audio spectrum signal received from a microphone to produce signals for stimulating a patient implantable tissue stimulating multi-channel electrode array adapted to be positioned in a cochlea from the apical region of the cochlea to the basal region of the cochlea, said method comprising selecting a first dominant frequency peak from said audio signal from a frequency band of between about 280Hz and about 1,000Hz and stimulating at least one electrode in the apical region of said electrode array in accordance with the spectral information contained in said first peak; selecting a second dominant frequency

peak from said audio signal from a frequency band of between about 800Hz and about 4,000Hz and stimulating at least one electrode in the basal region of said electrode array in accordance with the spectral information contained in said second peak; extracting spectral information in at least one region of the spectrum of said audio signal and stimulating at least one predetermined electrode in said electrode array in accordance with said extracted spectral information, said predetermined electrode being in said basal region of said electrode array.

Preferably, additional preselected electrodes are stimulated using spectral energy derived from said audio signal in the audio frequency regions 2,000 to 2,800Hz, 2,800 to 4,000Hz and above 4,000Hz respectively.

In accordance with another aspect of the invention there is provided an improved speech processor for a cochlear prosthesis which employs a multi-spectral peak (MPEAK) coding strategy to extract a number, for example five, of spectral peaks from an incoming acoustic signal received by a microphone. The speech processor encodes this information into sequential pulses that are sent to selected electrodes of a cochlear implant. The first formant (F1) spectral peak



(280-1000 Hz) and the second formant (F2) spectral peak (800-4000 Hz) are encoded and presented to apical and basal electrodes, respectively. F1 and F2 electrode selection follows the tonotopic organization of the cochlea. High-frequency spectral information is sent to more basal electrodes and low-frequency spectral information is sent to more apical electrodes. Spectral energy in the regions of 2000-2800 Hz, 2800-4000 Hz, and above 4000 Hz is encoded and presented to three fixed electrodes. The fundamental or voicing frequency (F0) determines the pulse rate of the stimulation during voiced periods and a pseudo-random aperiodic rate determines the pulse rate of stimulation during unvoiced periods. The amplitude of the acoustic signal in the five bands determines the stimulus intensity.

Brief Description of the Drawings

While the specification concludes with claims particularly pointing out and distinctly claiming the subject matter of this invention, it is believed that the invention will be better understood from the following description, taken in conjunction with the accompanying drawing, in which:

FIGS. 1A and 1B are interior views of the anatomy of a human ear and a cross section of a cochlea, respectively;

FIG. 2 is a block diagram of the overall cochlea implant system of this invention;

FIG. 3 is a pictorial view of the components of this system, including the implantable parts and the parts worn by the patient;

FIGS. 3A and 3B are respective side and end elevation views of the implantable parts of this system;

FIG. 4 is a graph of current vs. time, showing the biphasic current waveform utilized in this invention;

FIG. 5 is a graph showing an example of the sequential stimulation pattern of electrode pairs for a voiced sound using the multi-peak coding strategy of this invention;

FIG. 6 is a graph showing an example of the sequential stimulation pattern of electrode pairs for an unvoiced sound using the multi-peak coding strategy of this invention;

FIG. 7 is a chart showing an example of the pattern of electrical stimulation for various steady-state phonemes using the multi-peak coding strategy of this invention;

FIG. 8 is a graph showing the standard loudness growth function for the speech processor of this invention; and,

FIG. 9 is a block diagram of the microphone and speech processor portions of a pulsatile type, multi-channel cochlear implant system in accordance with this invention.

Best Mode for Carrying Out the Invention

The cochlear implant system of this invention, shown in FIG. 2, comprises several components. An electrode array 1 is implanted into the cochlea. The electrode array 1 comprises a number of rings or bands of platinum molded with a flexible silastic carrier. Preferably, there are 32 bands of platinum in total. The distal 22 bands are active electrodes, and have connecting wires welded to them. The proximal 10 electrode bands are used for stiffening, and to act as an aid to surgical insertion. In a typical array, the electrode rings are about 0.05 mm in thickness with a width of 0.3 mm, and have outside diame-

ters ranging from 0.6 mm at the proximal end to 0.4 mm diameter at the distal end. The diameter of the rings changes smoothly so that the array is tapered over the distal 10 mm or so. The rings are spaced on 0.75 mm centers over the distal 25 mm of the electrode array, and all of the exposed outside area of the rings is used as active electrode area. The silastic material may be MDX4-4210, manufactured by Dow Corning.

The 22 electrode wires pass via a cable 2 from the electrode array 1 to the receiver-stimulator unit (RSU) 3. The invention described is not limited to the use of this design of electrode array, and a number of alternative electrode designs as have been described in the prior art could be used. The RSU 3 receives information and power from an external source through a tuned receiving coil 5 attached to the RSU and positioned just beneath the skin. The RSU also provides electrical stimulating pulses to the electrode array 1. The power, and data on which electrode to stimulate, and with what intensity, is transmitted across the skin using an inductive link 6 operating at radio frequencies, from an external multipeak speech processor (MSP) 7. In normal operation, the MSP picks up acoustic stimuli from a microphone 8 conveniently worn, and extracts from the signal, information which is used to determine stimulation electrode, rate and amplitude.

Because each patient's response to electrical stimulation is different, it is necessary to configure each patient's MSP to his or her own requirements. Thus, the MSP has a random access memory (RAM) which is programmed to suit each patient.

The patient's response to electrical stimulation is tested some short time after implantation of the RSU 3, using the patient's MSP, and the results of

these tests are used to set up the MSP for the patient's own particular requirements. This is done by connecting the MSP, via a connector and cables 9, to a diagnostic programming interface unit (DPI) 10. The DPI is itself connected via a cable and connector 11 to a general purpose computer referred to as a diagnostic and programming unit (DPU) 12.

A pictorial representation of the system used by the patient is shown in FIGS. 3, 3A and 3B. The electrode array 20 is flexible and fits the shape of the cochlea (FIGS. 1A and 1B) as it is inserted along the basilar membrane separating the scala tympani from the remainder of the cochlea. The electrode array is connected via a silastic-covered cable 21 to the RSU 22. Cable 21 is specially designed to provide stress relief to prevent fracture of the wire in the cable. The receiving coil for information and power is a single turn of multistrand platinum wire 23 which is transformer coupled to the implanted electronics in the RSU 22.

An externally worn transmitting coil 24 is held against the head over the site of the RSU implant 22 by, for example, cooperating magnets (not shown) carried adjacent each of the coils 23 and 24, or by a fixture (not shown) attached to the coil 24 for holding the coil to the user's head, or by adhesive tape. Coil 24 is connected to the speech processor 29 via a coil cable 26 and a hearing aid microphone 27. Hearing aid microphone 27 is worn on the ear nearest to the implant site and audio data from the microphone 27 is connected via a three wire cable 28 to the MSP 29. Transmission data is connected to the coil 24 from the MSP 29 via the same three wire cable 28 and via the coil cable 26. This three wire arrangement is described in the copending U.S. Patent Application No.

404,230, filed September 7, 1989, of Christopher N. Daly, entitled "Three Wire System For Cochlear Implant Processor," which application is assigned to the assignee of the present invention and is incorporated herein by reference. Alternative microphone configurations are possible, including a microphone worn on a tie clasp or attached to the users clothing, or the like.

The coil cable 26 and three wire cable 28 are attached to the microphone 27 and MSP 29 by demountable connectors 32, 33 and 34. The MSP 29 is powered by conventionally available batteries (e.g., a single AA size cell) carried inside the MSP 29. A plug-in jack 31 is provided to allow connection of external audio signal sources, such as from a television, radio, or high quality microphone.

Referring to FIG. 4, the pulse which is used to electrically stimulate the cochlea is biphasic. That is, it comprises a period of negative current stimulation, followed by an equal period of positive current stimulation of equal amplitude, the two periods (known as phases $\phi 1$ and $\phi 2$), separated by a short period of no stimulation. $\phi 1$ and $\phi 2$ may be in the range of 12 to 400 microseconds (typically 200 microseconds), and the intervening interval is typically about 50 microseconds. The amplitudes of $\phi 1$ and $\phi 2$, their durations, and the duration of the intervening interval are determined by the information decoded from the signal transmitted by the speech processor 29 (FIG. 3). The actual values of these parameters will be set up on an electrode by electrode basis, for each patient, as a result of psychophysical testing of the patient. The reversal in polarity of $\phi 1$ and $\phi 2$ is important since it ensures that there is no net DC component in the stim-

ulus. This is important because long term DC excitation might cause electrode corrosion, and possible subsequent damage to the cochlea itself.

The questions of electrode electrochemistry and charge balance are thought to be more important in cochlear implants than in, say, cardiac pacemakers which are well known in the art. This is because a cochlear stimulator will be stimulating nerve fibers, whereas a cardiac pacemaker is designed to stimulate cardiac muscle. It is thought that nerve tissues may be more susceptible to damage due to electrical stimulation, and thus the cochlear implant system is designed with more stringent safety factors than cardiac pacemakers. The system is designed so that the same stimulus source is used for both stimulation phases. The biphasic pulse is produced simply by reversal of the connections to the electrodes. Thus, extremely good charge symmetry is obtained, resulting in a high level of safety, provided the durations of ϕ_1 and ϕ_2 are equal.

The stimulation circuitry is preferably configured as a constant current source. This has the advantage compared to a constant voltage source that if the electrode impedance changes (as has often been observed) the delivered current to the electrode will remain unaltered over a large range of electrode impedances. The current may be varied from a few microamps to 2 mA, allowing a very large range of loudness percepts to be produced and large variations between patients to be accommodated.

The stimulus generation circuitry in the RSU 3 (FIG 2) is preferably designed to operate in one of two modes. The first mode is referred to as "multipolar" or "common ground" stimulation. In this mode, one electrode is selected to be the "active" elec-

trode, and all other electrodes operate as a common current source. In phase phi 2, the connections are reversed so that the "active" electrode acts as the current source and the common electrodes act as a current sink. The choice of stimulus order is not determined by any limitations or restrictions in the circuit design, and either way may be chosen when implementing the circuit design.

The second mode is "bipolar" stimulation. In this mode stimulation is between two selected electrodes, let us say A and B. In phase phi 1, current is sourced by A, and sunk by B. In phase phi 2, current is sourced by B, and sunk by A, and no other electrodes play any part in stimulation. The RSU 3 is preferably configured so that any pair of electrodes may be selected for bipolar stimulation. Thus there is great flexibility in choice of stimulation strategy.

It should be understood that only these two particular stimulation modes have been chosen. Other stimulation modes are not excluded, however. For example, a multipolar or distributed ground system could be used where not all other electrodes act as a distributed ground, and any electrode could be selected at any time to be a current source, current sink, or inactive during either stimulation phase with suitable modification of the receiver-stimulator.

The main aim of this invention is to provide improved speech communication to those people suffering from profound hearing loss. However, in addition to providing improved speech communication, it is also important to be able to convey environmental sounds, for example telephones, doors, warning sirens, door bells, etc., which form part of a person's life. The system described up to now is basically that of the

Crosby et al. patent, heretofore referred to and incorporated herein by reference. In the Crosby et al. patent it is recognized that the second formant F2 carries most of the intelligibility of the speech signal, while the first formant F1, although containing much of the naturalness of the signal, contributes little to intelligibility.

Crosby et al. observed that the third and higher formants do not carry as much information as the second formant. They also felt that in view of the then limitations of knowledge on the interaction between electrodes when a number of electrodes are stimulated simultaneously, the most effective method of stimulation would be to code the second formant on an appropriate electrode or site in the cochlea to provide the most important formant information. The amplitude of such stimulation is derived from the amplitude of the second formant.

The Crosby et al. system also provides prosodic information in the form of pulse rate. That system compresses the stimulation rate to the range 100 - 250 Hz, the range in which the greatest pitch discrimination from stimulation pulse rate is achieved.

An additional factor employed in Crosby et al. is that only the top 10 to 20 dB of current acoustic stimulus level is used to determine stimulus amplitude. That is, instead of compressing the entire acoustic loudness range into the small range of electrical stimulation available, only the top part is used. Thus, Crosby et al.'s amplitude of the signal is entirely represented by a five bit binary code, which provides only 30 dB of dynamic range.

In summary, the Crosby et al. speech processing strategy is:

1. The dominant spectral peak in the range of about 300 Hz to about 4000 Hz is used to encode electrode position.

2. The amplitude of the dominant spectral peak used to encode electrode position is used to determine stimulation amplitude.

3. Voice pitch (F0) is compressed and used to determine the stimulation rate.

For unvoiced sounds and, environmental sounds, the Crosby et al. system still generates stimuli, but the stimulation rate and electrode position will be determined by the exact nature of the acoustic signal. For example, for sibilant consonants ("s"), the stimulation rate is fairly fast, but not constant, and the electrode stimulated will be one which elicits a high frequency percept.

A second speech processing strategy, useful in some patients, is employed in Crosby et al. The second strategy is similar to the one mentioned above in that electrode position is encoded from formant frequency. However, the stimulation rate is at the F1 of first formant frequency, and the stimulation amplitude is determined for the value of the peak of the acoustic signal at the time of the F1 peak. This has the advantage that the stimulation rate is faster, and elicits more natural sounding speech perceptions in some patients. In addition, since the F1 signal is amplitude modulated and temporally better than the F0 rate, the patients also perceive the F0 or voice pitch which is useful for conveying prosodic information.

Another speech processing strategy considered in the Crosby et al. reference is to stimulate the patient at the rate of F1 extracted from an incoming speech signal, but to pattern the stimulation such that the stimuli are gated at the F0 rate.

Notwithstanding the success of speech processors using the Crosby et al. F0, F1, F2 speech processing coding scheme over the last few years, a number of problems still remain in connection with the use of such speech processor coding schemes. As indicated earlier, patients who perform well in quiet conditions can have significant problems when there is a moderate level of background noise. Moreover, since the F0, F1, F2 scheme codes frequencies up to about 4000Hz, and many phonemes and environmental sounds have a high proportion of their energy above this range, such phonemes and environmental sounds are inaudible to the implant user in some cases.

In accordance with the present invention multichannel cochlear implant prostheses having a pulsatile operating system, such as that disclosed in the Crosby et al. reference, are provided with a speech coding scheme in which the speech signal is bandpass filtered into a number of bands, for example 3, within and beyond the normal range of the second frequency peak or formant F2 of the speech signal. The speech coding scheme disclosed herein is referred to as the multi-spectral peak coding strategy (MPEAK). MPEAK is designed to provide additional high-frequency information to aid in the perception of speech and environmental sounds.

The MPEAK coding strategy extracts and codes the F1 and F2 spectral peaks, using the extracted frequency estimates to select a more apical and a more basal pair of electrodes for stimulation. Each selected electrode is stimulated at a pulse rate equal to the fundamental frequency F0. In addition to F1 and F2, three high frequency bands of spectral information are extracted. The amplitude estimates from band three (2000-2800 Hz), band four (2800-4000 Hz),

and band five (above 4000 Hz) are presented to fixed electrodes, for example the seventh, fourth and first electrodes, respectively, of the electrode array 1 (FIG. 2).

The first, fourth and seventh electrodes are selected as the default electrodes for the high-frequency bands because they are spaced far enough apart so that most patients will be able to discriminate between stimulation at these three locations. Note that these default assignments may be reprogrammed as required. If the three high frequency bands were assigned only to the three most basal electrodes in the MAP, many patients might not find the additional high frequency information as useful since patients often do not demonstrate good place-pitch discrimination between adjacent basal electrodes. Additionally, the overall pitch percept resulting from the electrical stimulation might be too high.

Table I below indicates the frequency ranges of the various formants employed in the speech coding scheme of the present invention.

TABLE I

<u>Frequency Range</u>	<u>Formant or Band</u>
280 - 1000 Hz	F1
800 - 4000 Hz	F2
2000 - 2800 Hz	Band 3 - Electrode 7
2800 - 4000 Hz	Band 4 - Electrode 1
4000 Hz and above	Band 5 - Electrode 1

If the input signal is voiced, it has a periodic fundamental frequency. The electrode pairs selected from the estimates of F1, F2 and bands 3 and 4 are stimulated sequentially at the rate equal to F0. The most basal electrode pair is stimulated first, followed by progressively more apical electrode pairs, as shown in FIG. 5. Band 5 is not presented in FIG. 5

because negligible information is contained in this frequency band for voiced sounds.

If the input energy is unvoiced, energy in the F1 band (280-1000 Hz) is effectively zero. Consequently it is replaced with the frequency band that extracts information above 4000 Hz. In this situation, the electrodes pairs selected from the estimates of F2, and bands 3, 4 and 5 receive the pulsatile stimulation. The rate of stimulation is aperiodic and varies between 200-300 Hz. FIG. 6 shows the sequential stimulation pattern for an unvoiced sound, with stimulation progressing from base to apex. The MPEAK coding strategy thus may be seen to extract and code five spectral peaks but only four spectral peaks are encoded for any one stimulus sequence.

FIG. 7 illustrates the pattern of electrical stimulation for various steady state phonemes when using the MPEAK coding strategy. A primary function of the MAP is to translate the frequency of the dominant spectral peaks (F1 and F2) to electrode selection. To perform this function, the electrodes are numbered sequentially starting at the round window of the cochlea. Electrode 1 is the most basal electrode and electrode 22 is the most apical in the electrode array. Stimulation of different electrodes normally results in pitch perceptions that reflect the tonotopic organization of the cochlea. Electrode 22 elicits the lowest place-pitch percept, or the "dullest" sound. Electrode 1 elicits the highest place-pitch percept, or "sharpest" sound.

To allocate the frequency range for the F1 and F2 spectral peaks to the total number of electrodes, a default mapping algorithm splits up the total number of electrodes available to use into a ratio of approximately 1:2, as shown in FIG. 7. Con-

sequently, approximately one third of the electrodes are assigned to the F1 frequency range. These are the more apical electrodes and they will cover the frequency range of 280-1000 Hz. The remaining two thirds of the electrodes are assigned to the F2 frequency range (800-4000 Hz). The most apical electrodes, which cover the frequency range from 280-1000 Hz, are assigned linearly equal frequency bands. The frequency range corresponding to the estimate of F2 is assigned to the remaining more basal electrodes and is divided into logarithmically equal frequency bands. This frequency distribution is called linear/log (lin/log) spacing.

A second optional mapping algorithm (not shown) splits up the total frequency range into logarithmically equal frequency bands for both F1 and F2 electrode groups (log/log spacing). In comparison to the lin/log spacing, this results in relatively broad frequency bands for electrodes that are assigned frequency boundaries below 1000 Hz. Because of the wider frequency bands for these electrodes, many vowel sounds will stimulate similar electrodes, thus making discrimination of these vowels difficult.

The F1/F2 lin/log function of the default algorithm is preferable because it gives better spatial resolution in the F1 range than the log/log function. In addition, this algorithm provides discrimination of vowels and consonants with formants close to 1000 Hz.

The mapping section of the DPS program allows flexibility in assigning frequency bands to electrodes. If fewer electrodes are included in the MAP, then fewer and wider frequency bands are allocated automatically by the computer so that the entire frequency range is covered. Furthermore, it is possible

to override the computer-generated spacing of frequency bands. Any range of frequencies may be allocated to any electrode or electrodes by changing the upper frequency boundaries.

Table II, below, shows the default boundaries (lin/log) for a MAP created in the biphasic +1 mode using 20 electrode pairs and the MPEAK coding strategy.

TABLE II - Lin/Log Frequency Boundaries for 20 Electrodes in a BP +1 Mode. Also Shown are the electrode allocations for the three high frequency bands.

Electrode	<u>Frequency Boundaries</u>	
	Lower	Upper
20	280	400
19	400	500
18	500	600
17	600	700
16	700	800
15	800	900
14	900	1000
13	1000	1112
12	1112	1237
11	1237	1377
10	1377	1531
9	1531	1704
8	1704	1896
7	1896	2109
6	2109	2346
5	2346	2611
4	2611	2904
3	2904	3231
2	3231	3595
1	3595 & above	

Electrodes: for Band 3 - 7
for Band 4 - 4
for Band 5 - 1

Table III, below, shows the default boundaries in the same mode using only 14 electrode pairs and the MPEAK coding strategy.

TABLE III - Lin/Log Frequency Boundaries for 14 Electrodes in a BP +1 Mode. Also

shown are the electrode allocations
for the three high frequency bands.

Electrode	<u>Frequency Boundaries</u>	
	Lower	Upper
20	280	400
18	400	550
17	550	700
16	700	850
15	850	1000
14	1000	1166
13	1166	1360
10	1360	1587
9	1587	1851
8	1851	2160
7	2160	2519
6	2519	2939
5	2939	3428
4	3428 & above	

Electrodes: for Band 3 - 8
for Band 4 - 6
for Band 5 - 4

The amplitude of the electrical stimulus is determined from the amplitude of the incoming acoustic signal within each of the five frequency bands (F1, F2, Bands 3, 4 and 5). However, because the electrodes have different threshold (T) and maximum acceptable loudness (C) levels, the speech processor must determine the level of stimulation for each electrode separately based on the amplitude of the incoming signal in each band.

The MSP (FIG. 2) contains a non-linear loudness growth algorithm that converts acoustic signal amplitude to electrical stimulation parameters. First, the MSP converts the amplitude of the acoustic signal into a digital linear scale with values from 0 to 150, as may be seen now by reference to FIG. 8. That digital scale (in combination with the T and C-levels stored in the patient's MAP) determines the actual charge delivered to the electrodes. Signals

whose amplitude levels are coded as 1 will cause stimulation at the T-level. Signals whose amplitude levels are coded as 150 will cause stimulation at the C-level.

Referring now to FIG. 9, a block diagram of the microphone and speech processor portions of a pulsatile type, multi-channel cochlear implant system 100 have there been illustrated. The system 100 includes a microphone 110 which picks up speech and provides electrical audio signals to a speech feature extractor 112 through an automatic gain control amplifier 111. The speech feature extractor 112 analyzes the signals and provides digital outputs corresponding to the frequencies and amplitudes of the first and second formants, identified as F1, A1, F2 and A2, respectively, in FIG. 10.

The speech feature extractor 112 also detects and outputs the voice pitch FO and starts the encoder 113, which translates, using a MAP 114 containing information on the patient's psychophysical test results, the voice pitch information into a pattern of electrical stimulation on two electrodes that are stimulated sequentially. The data so translated is sent by the patient coil 115 to the implanted receiver stimulator unit RSU 3 (FIG. 2).

Three bandpass filters 116, 117 and 118 also receive the audio signal from microphone 110 before it is applied to the speech feature extractor 112, and separate the signal into three components of different frequencies, a 2000-2800 Hz signal in band 3, a 2800-4000 Hz signal in band 4, and 4000-8000 Hz signal in band 5. The signals from bands 3, 4 and 5 are lead to the encoder 113 and mapping of these signals is done in a manner similar to that for the first and second formants, and translation of the resulting pattern of

electrical stimulation to the appropriate electrodes takes place, as discussed earlier herein.

The automatic gain control amplifier 111 is used to control the amplitude of the signal fed to the filters 116 and 117. Since filter 118 is only used for unvoiced parts of the speech signal, its amplitude is never very great and, therefore, the signal does not require automatic gain control. Accordingly, amplifier 119 does not have automatic gain control provisions incorporated therein.

To summarize, the psychophysical measurements that are made using the DPS software provide the information for translating the extracted acoustic input into patient-specific stimulation parameters. Threshold (T) and maximum (C) levels for electrical stimulation are measured for each electrode pair. These values are stored in the MAP. They determine the relationship between the incoming acoustic signal amplitude and the stimulation level for any given electrode pair.

Inside the speech processor a random access memory stores a set of number tables, referred to collectively as a MAP. The MAP determines both stimulus parameters for F1, F2 and bands 3-5, and the amplitude estimates. The encoding of the stimulus parameters follows a sequence of distinct steps. The steps may be summarized as follows:

1. The first formant frequency (F1) is converted to a number based on the dominant spectral peak in the region between 280-1000 Hz.
2. The F1 number is used, in conjunction with one of the MAP tables, to determine the electrode to be stimulated to represent the first formant. The indifferent electrode is determined by the mode.

3. The second formant frequency (F2) is converted to a number based on the dominant spectral peak in region between 800-4000 Hz.

4. The F2 number is used, in conjunction with one of the MAP tables to determine the electrode to be stimulated to represent the second formant. The indifferent electrode is determined by the mode.

5. The amplitude estimates for bands 3, 4 and 5 are assigned to the three default electrodes 7, 4 and 1 for bands 3, 4 and 5, respectively, or such other electrodes that may be selected when the MAP is being prepared.

6. The amplitude of the acoustic signal in each of the frequency bands is converted to a number ranging from 0 - 150. The level of stimulation that will be delivered is determined by referring to a set MAP tables that relate acoustic amplitude (in range of 0-150) to stimulation level for the specific electrodes selected in steps 2, 4 and 5, above.

7. The data are further encoded in the speech processor and transmitted to the receiver/stimulator. It, in turn, encodes the data and sends the stimuli to the appropriate electrodes. Stimulus pulses are presented at a rate equal to F0 during voiced periods and at a random aperiodic rate within the range of F0 and F1 formants (typically 200 to 300 Hz) during unvoiced periods.

It will be apparent from the foregoing description that the multi-spectral peak speech coding scheme of the present invention provides all of the information available in the prior art F0F1F2 scheme, while providing additional information from three high frequency band pass filters. These filters cover the following frequency ranges: 2000 to 2800 Hz, 2800 to 4000 Hz and 4000 to 8000 Hz. The energy within these

ranges controls the amplitude of electrical stimulation of three fixed electrode pairs in the basal end of the electrode array. Thus, additional information about high frequency sounds is presented at a tonotopically appropriate place within the cochlea.

The overall stimulation rate remains as F0 (fundamental frequency or voice pitch) but in the scheme of the present invention four electrical stimulation pulses occur for each glottal pulse. This compares with the prior F0F1F2 strategy in which only two pulses occur per voice pitch period. In the new coding scheme, for voiced speech sounds, the two pulses representing the first and second formant are still provided, and additional stimulation pulses occur representing energy in the 2000-2800 Hz and the 2800-4000 Hz ranges.

For unvoiced phonemes, yet another pulse representing energy above 4000 Hz is provided while no stimulation for the first formant is provided, since there is no energy in this frequency range. Stimulation occurs at a random pulse rate of approximately 260 Hz, which is about double that used in the earlier strategy.

It will be further apparent from the foregoing description that this invention provides an improved cochlear implant system which overcomes various of the problems associated with earlier cochlear implant systems. The use of a multi-spectral peak speech coding strategy in accordance with this invention provides the user of the implant system with significantly improved speech recognition, even in the presence of moderate levels of background noise. In addition improved recognition of phonemes and environmental sounds are provided by this invention.

While a particular embodiment of this invention has been shown and described, it will be obvious to those skilled in the art that various changes and modifications may be without departing from this invention in its broader aspects, and it is, therefore, aimed in the appended claims to cover all such changes and modifications as fall within the true spirit and scope of this invention.

Claims

What is claimed is:

1. A multi-channel cochlear prosthesis including a patient implantable tissue stimulating multi-channel electrode array adapted to be positioned in a cochlea from the apical region of the cochlea to the basal region of the cochlea, a patient implantable multi-channel stimulator connected to said array, and a patient externally worn programmable speech processor for processing sound signals into electrical stimulation signals that are transmitted to said stimulator, said prosthesis further comprising:

means based on a dominant peak extraction in the region of between about 280 Hz to about 1000 Hz for determining a first formant spectral information in said sound signal and stimulating at least one electrode in the apical region of said electrode array in accordance with the spectral information of said formant;

means based on a dominant spectral peak extraction in the region of between about 800 Hz and about 4000 Hz for determining a second formant frequency in said sound signal and stimulating at least one electrode in said basal region of said electrode array in accordance with the spectral information of said formant; and,

at least one high frequency band filter for extracting spectral information in at least one region of the spectrum of said sound signal and stimulating at least one predetermined electrode in said electrode array in accordance with said extracted spectral information, said predetermined electrode being in said basal region of said electrode array.

2. A multi-channel cochlear prosthesis according to claim 1, including a plurality of said

high frequency band filters for extracting spectral information in a corresponding number of regions of said sound signal and stimulating at least a corresponding number of said predetermined electrodes in said electrode array, all of said predetermined electrodes being in said basal region of said electrode array.

3. A multi-channel cochlear prosthesis according to claim 2, wherein said electrical stimulation signals are applied to said electrodes in the form of pulses presented at a pulse rate dependent on the pitch of the sound signal.

4. A multi-channel cochlear prosthesis according to claim 3, wherein said pulse rate is in the range of between about 80 Hz and about 400 Hz.

5. A multi-channel cochlear prosthesis according to any one of claims 1 - 4 including at least three of said high frequency band filters each with predetermined electrodes.

6. A multi-channel cochlear prosthesis according to claim 5, wherein a first one of said high frequency band filters extracts spectral information from the sound signal in a frequency range of between about 2000 Hz and about 2800 Hz, wherein a second one of said high frequency band filters extracts spectral information from the sound signal in a frequency range of between about 2000 Hz and about 2800 Hz, and wherein a third one of said high frequency band filters extracts spectral information from the sound signal in a frequency range of between about 4000 Hz and about 8000 Hz.

7. A multi-channel cochlear prosthesis according to claim 6, wherein the electrodes in said electrode array may be considered to be consecutively numbered, starting from the basal end thereof and extending to the apical end thereof, and wherein am-

plitude estimates derived from the spectral information extracted from said first, second and third high frequency band filters are applied to said corresponding number of predetermined electrodes with the amplitude estimate from said first filter being applied to a higher-numbered electrode than the amplitude estimate from said second filter and the amplitude estimate from said second filter being applied to a higher-numbered electrode than the amplitude estimate from said third filter.

8. A multi-channel cochlear prosthesis according to claim 7, wherein said electrode array includes about 22 electrodes therein, wherein said basal region of said electrode array comprises about two-thirds of the electrodes in said electrode array, and wherein said apical region comprises about one-third of the electrodes in said electrode array.

9. A multi-channel cochlear prosthesis according to claim 8, wherein said amplitude estimates derived from the spectral information extracted from said first, second and third high frequency band filters are applied to said seventh, fourth and first electrodes, respectively, in said electrode array.

10. A multi-channel cochlear prosthesis according to claim 7, wherein, in the case of voiced sound signals, the electrodes selected to be stimulated are based on the first and second formants and on information derived from the first and second filters, and wherein said electrodes are stimulated sequentially at a rate that is based on the pitch of the sound signal, with the most basal of said electrodes being stimulated first, followed by stimulation of progressively more apical electrodes.

11. A multi-channel cochlear prosthesis according to claim 7, wherein, in the case of unvoiced

sound signals, the electrodes selected to be stimulated are based on the second formant and on information derived from the first, second and third filters, and wherein said electrodes are stimulated sequentially at an aperiodic rate within the range of from formant F0 to formant F1, with the most basal of said electrodes being stimulated first, followed by stimulation of progressively more apical electrodes.

12. A multi-channel cochlear prosthesis according to claim 7, wherein said electrodes are stimulated at an aperiodic rate within the range of about 200Hz to about 300Hz.

13. A multi-channel cochlear prosthesis according to claim 11, wherein said aperiodic rate is within the range of about 200Hz to about 300Hz.

14. A method of processing an audio spectrum signal received from a microphone to produce signals for stimulating a patient implantable tissue stimulating multi-channel electrode array adapted to be positioned in a cochlea from the apical region of the cochlea to the basal region of the cochlea, said method comprising selecting a first dominant frequency peak from said audio signal from a frequency band of between about 280Hz and about 1,000Hz and stimulating at least one electrode in the apical region of said electrode array in accordance with the spectral information contained in said first peak; selecting a second dominant frequency peak from said audio signal from a frequency band of between

about 800Hz and about 4,000Hz and stimulating at least one electrode in the basal region of said electrode array in accordance with the spectral information contained in said second peak; extracting spectral information in at least one region of the spectrum of said audio signal and stimulating at least one predetermined electrode in said electrode array in accordance with said extracted spectral information, said predetermined electrode being in said basal region of said electrode array.

15. A method of processing an audio spectrum signal as claimed in claim 14 wherein additional preselected electrodes are stimulated using spectral energy derived from said audio signal in the audio frequency regions 2,000 to 2,800Hz, 2,800 to 4,000Hz and above 4,000Hz respectively.

AMENDED CLAIMS

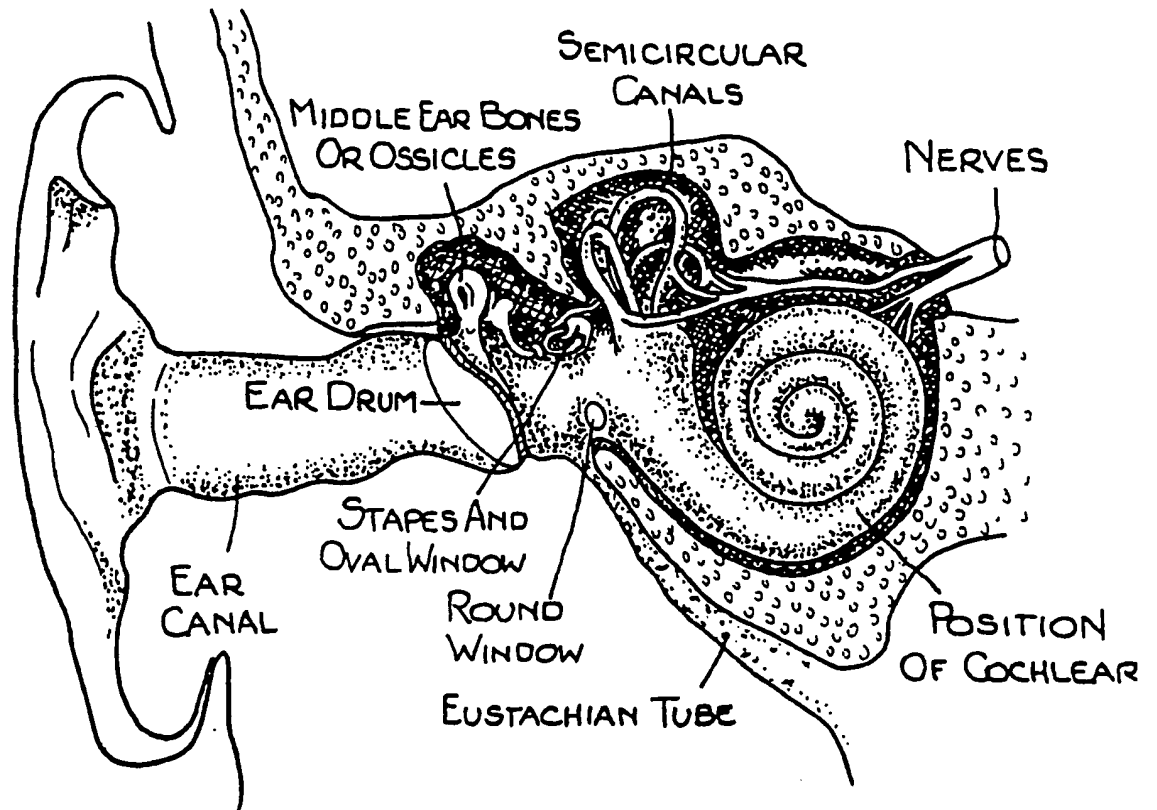
[received by the International Bureau on
14 January 1991 (14.01.91);
new claims 16 and 17 added;
other claims unchanged (1 page)]

about 800Hz and about 4,000Hz and stimulating at least one electrode in the basal region of said electrode array in accordance with the spectral information contained in said second peak; extracting spectral information in at least one region of the spectrum of said audio signal and stimulating at least one predetermined electrode in said electrode array in accordance with said extracted spectral information, said predetermined electrode being in said basal region of said electrode array.

15. A method of processing an audio spectrum signal as claimed in claim 14 wherein additional preselected electrodes are stimulated using spectral energy derived from said audio signal in the audio frequency regions 2,000 to 2,800Hz, 2,800 to 4,000Hz and above 4,000Hz respectively.

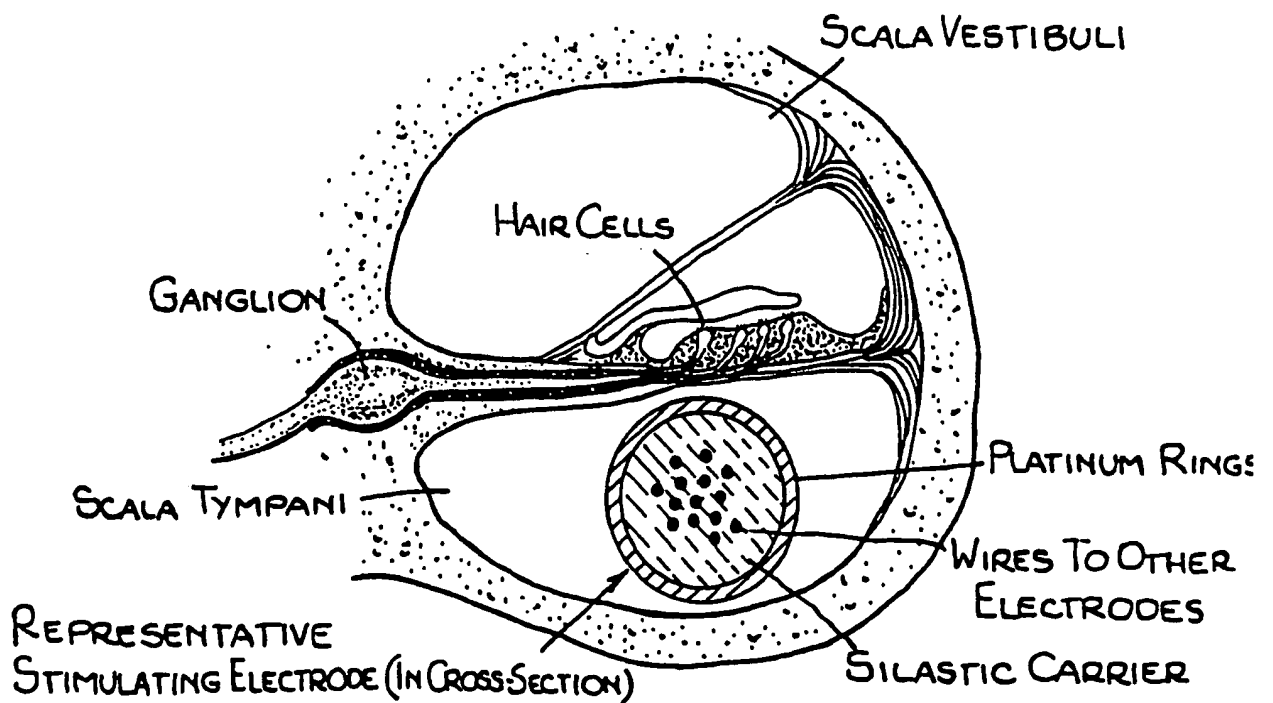
16.(new) A method of speech coding an audio spectrum signal received from a microphone to produce signals for stimulating a patient implantable tissue stimulating multi-channel electrode array adapted to be positioned in a cochlea from the apical region of the cochlea to the basal region of the cochlea, said method comprising bandpass filtering said audio spectrum signal into a plurality of bands within and beyond the normal range of a second (F2) formant frequency peak of said audio spectrum signal whereby additional high frequency information is provided to a patient.

17.(new) The method of claim 16 wherein information thus derived from said audio spectrum signal is encoded into sequential pulses and applied to selected electrodes of said electrode array.



ANATOMICAL ARRANGEMENT

Fig. 1A.

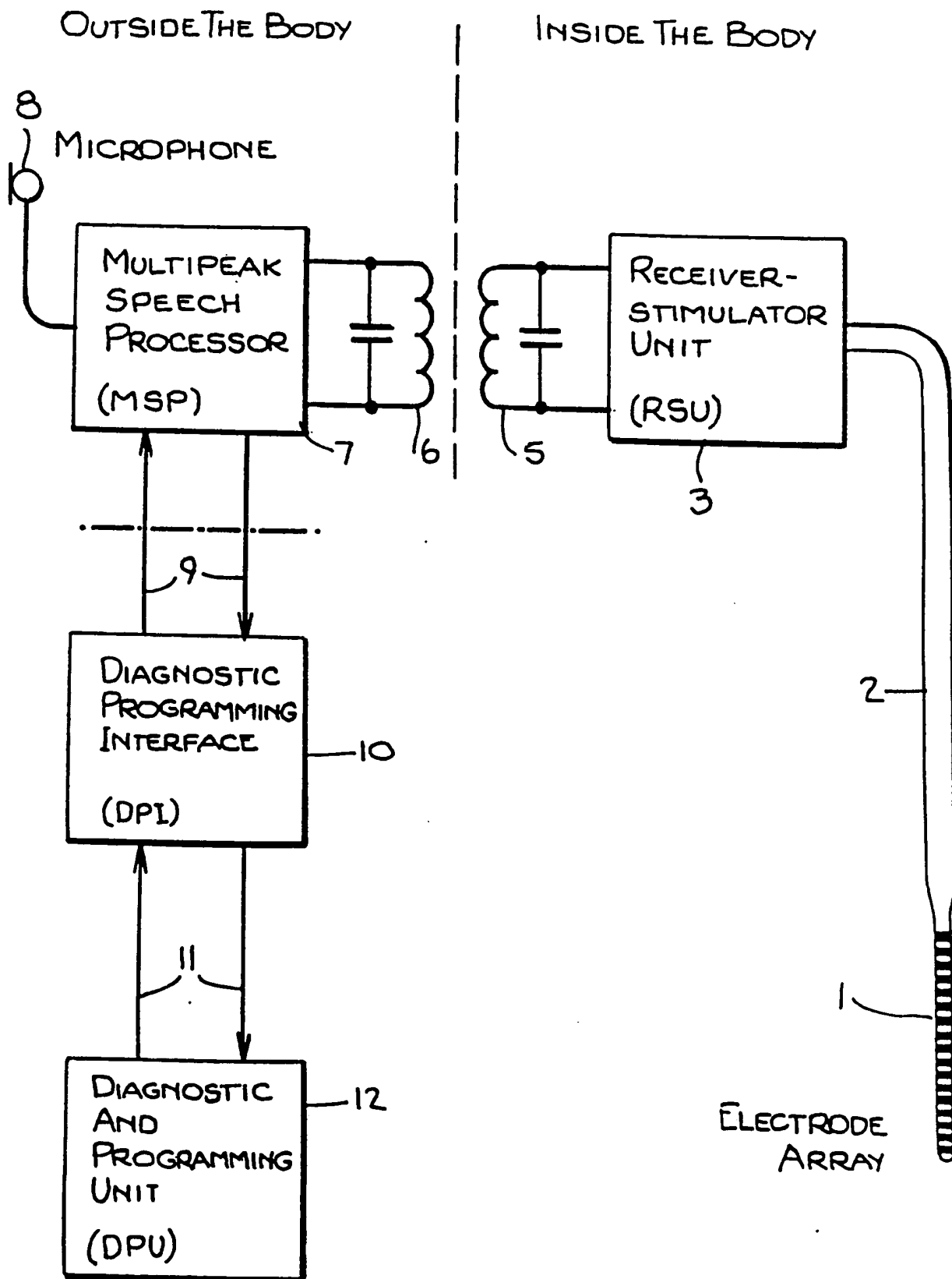


COCHLEA IN CROSS-SECTION

Fig. 1B.

2/7 Fig. 2.

COCHLEAR IMPLANT SYSTEM BLOCK DIAGRAM



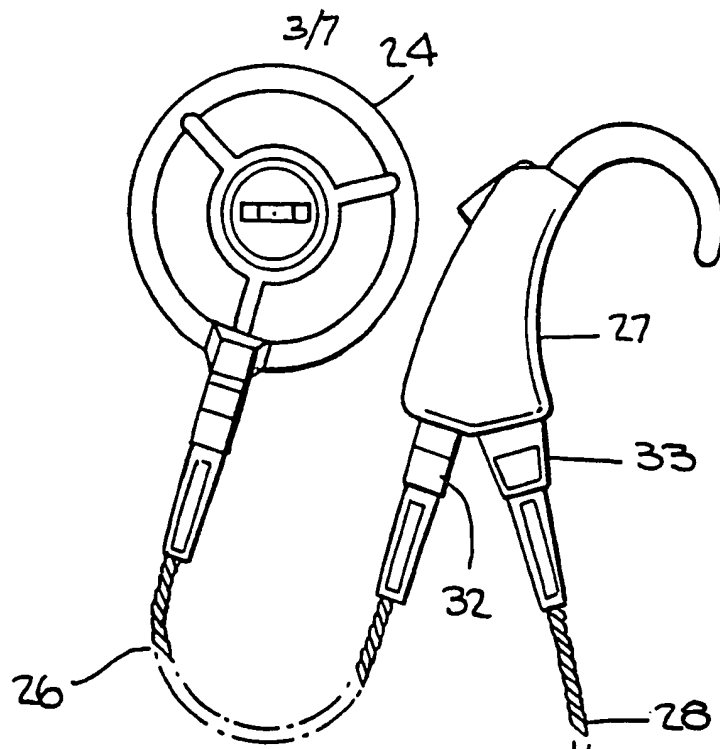


Fig. 3.

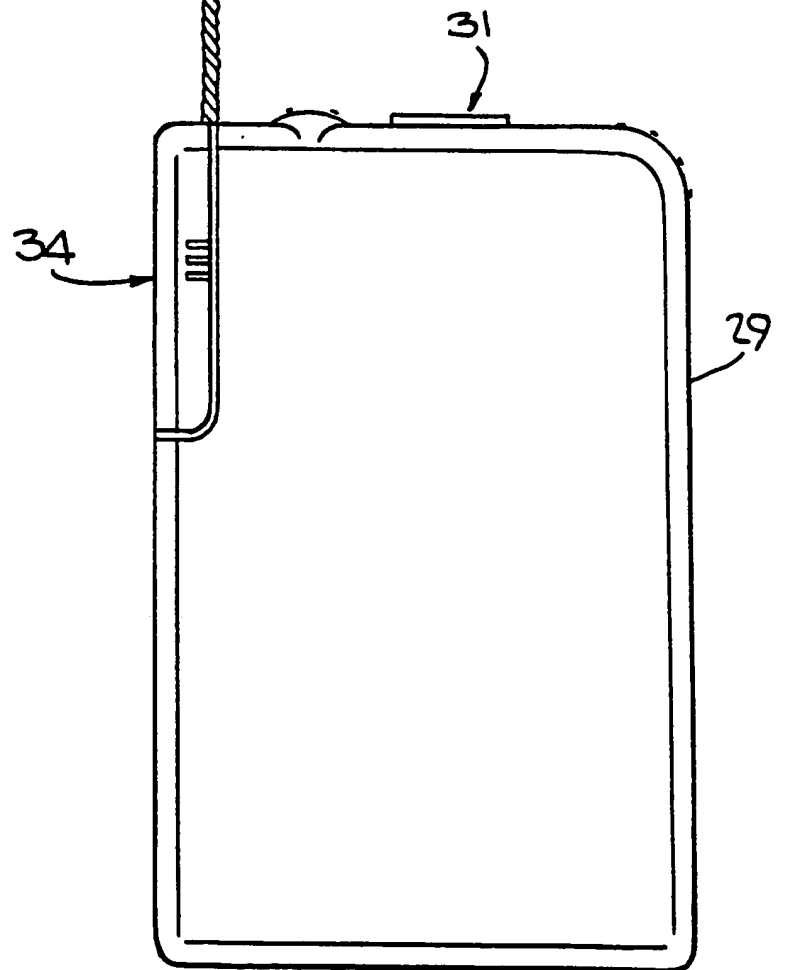


Fig. 3A.

4/7

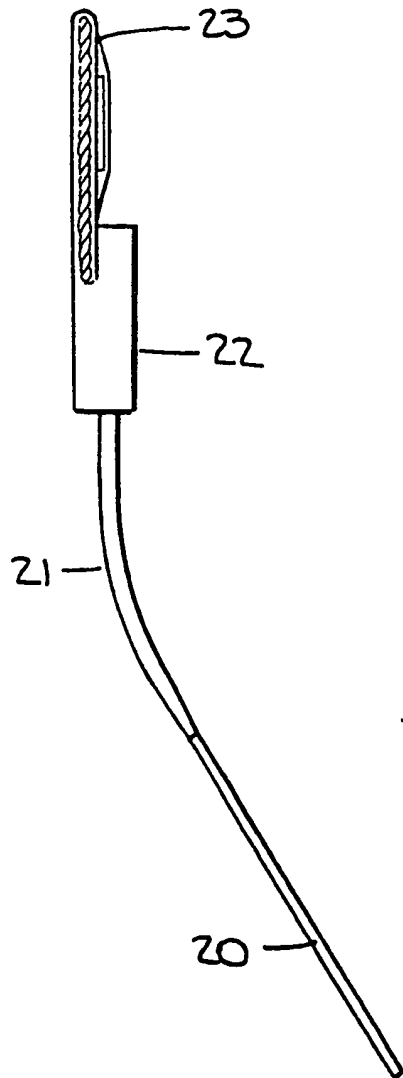
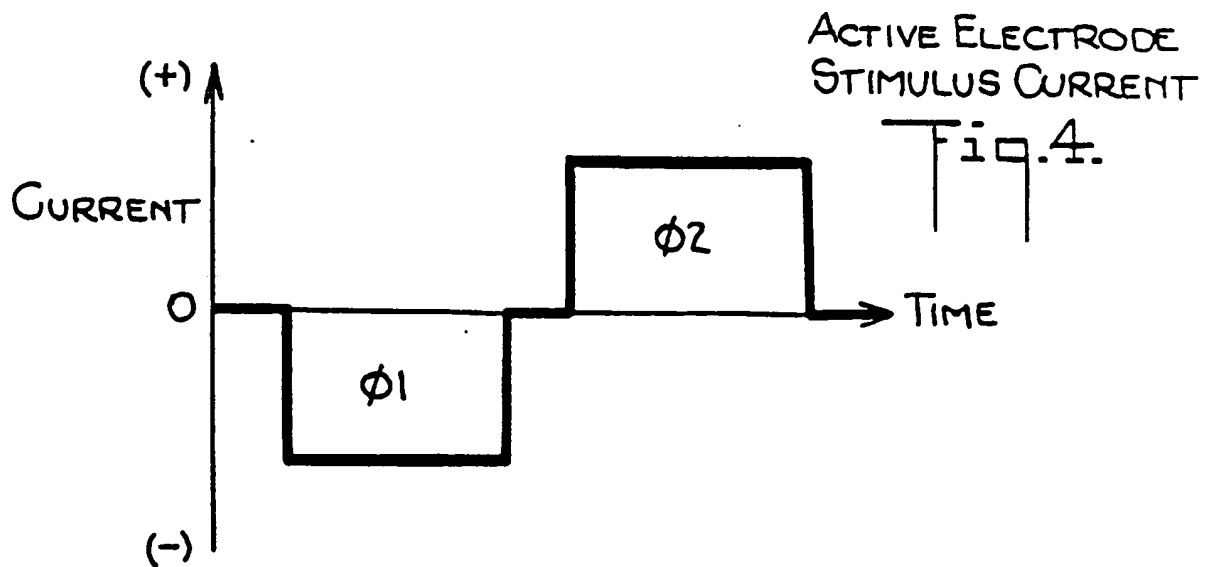
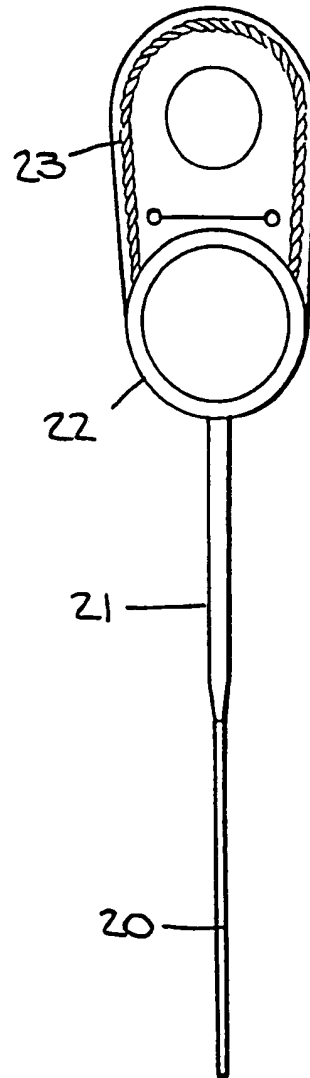
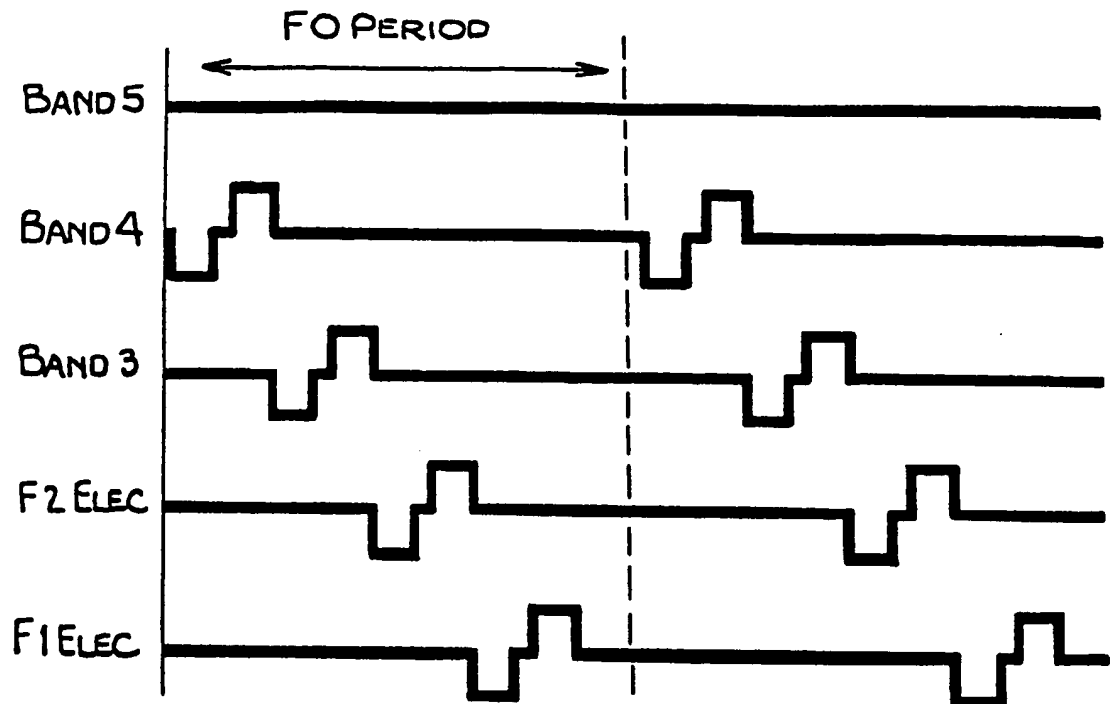


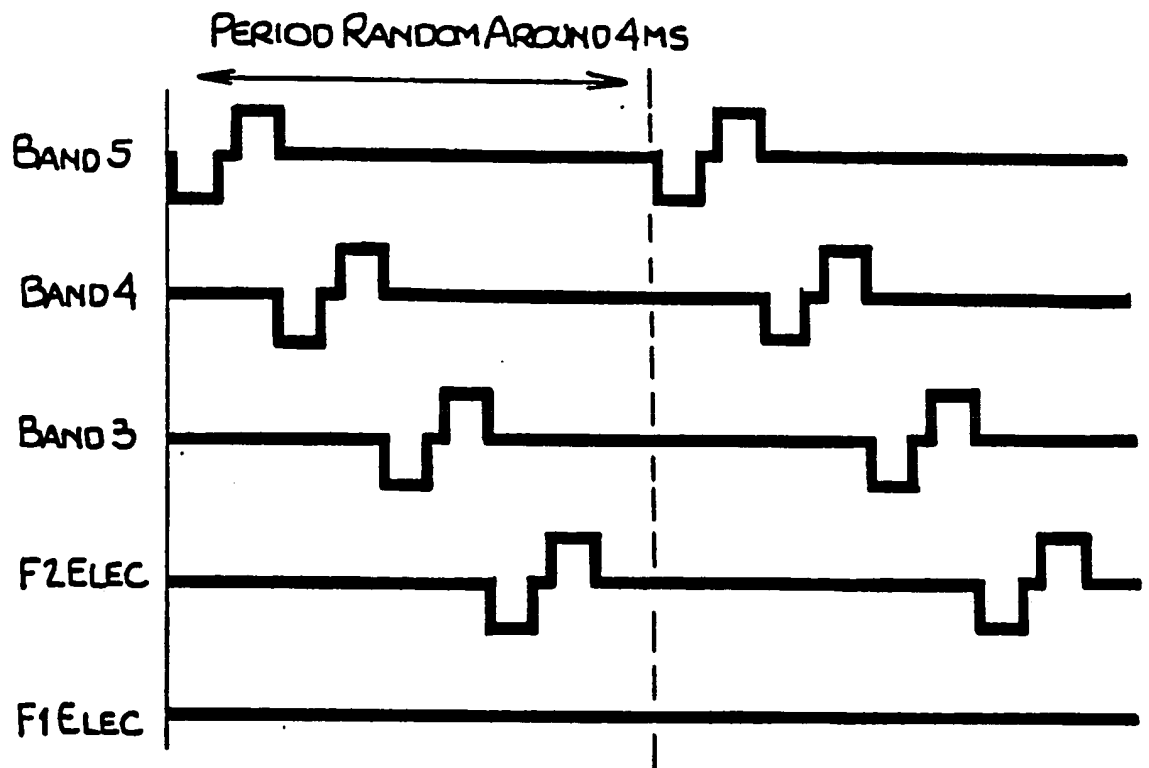
Fig. 3B.





MPEAK STRATEGY-VOICED SOUNDS

Fig. 5.

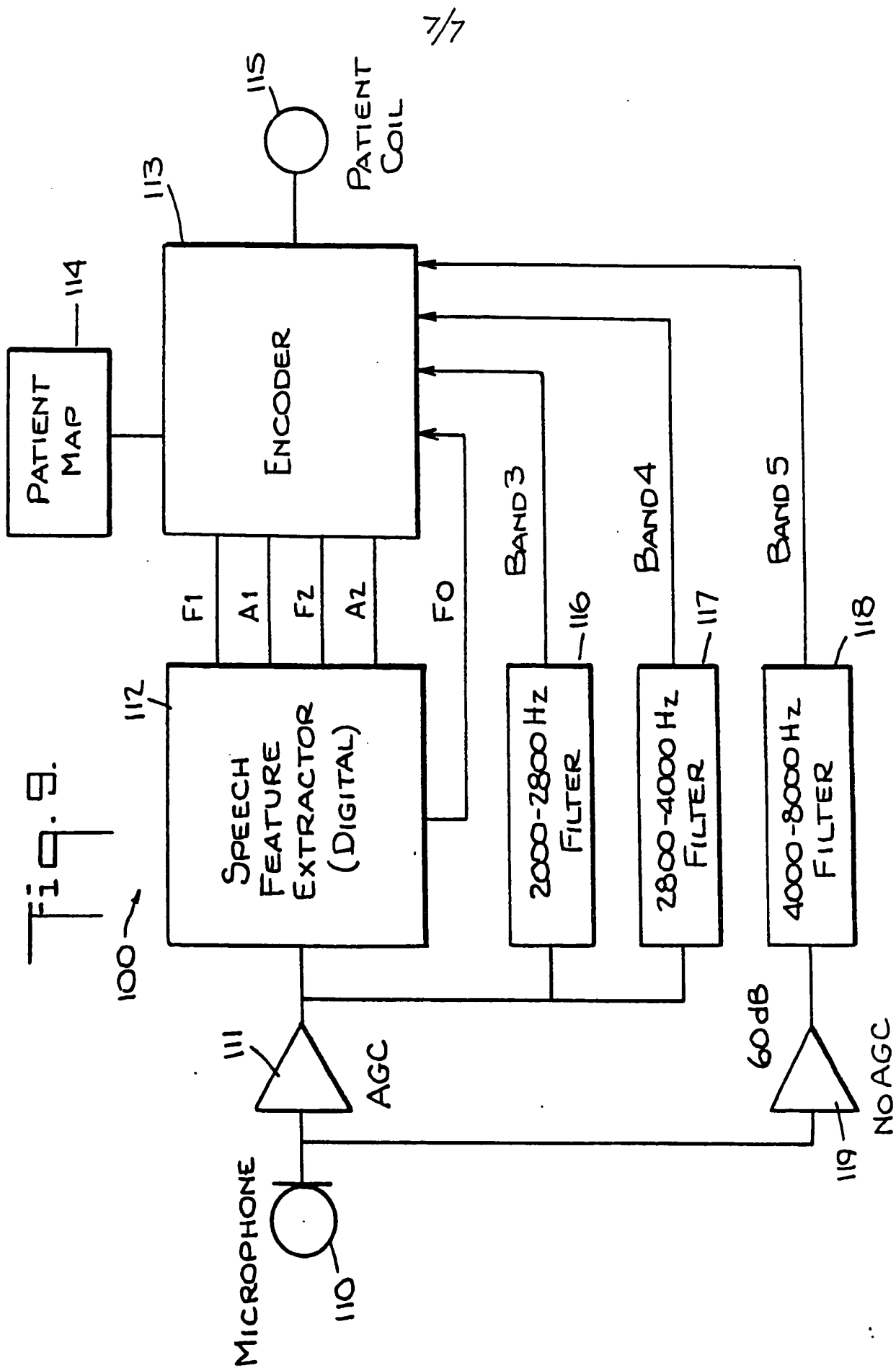


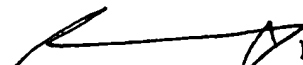
MPEAK STRATEGY-UNVOICED SOUNDS

Fig. 6.

/s/ /z/ /a/ /u/ /i/ BASAL (HIGH PITCH)

Fig. 9.



I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) 6		
According to International Patent Classification (IPC) or to both National Classification and IPC		
Int. Cl. ⁵ H04R 25/00, A61F 2/18		
II. FIELDS SEARCHED		
Minimum Documentation Searched 7		
Classification System	Classification Symbols	
I.	H04R 25/00, A61F 2/18, 1/18	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched 8		
AU : IPC as above		
III. DOCUMENTS CONSIDERED TO BE RELEVANT 9		
Category*	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages 12	Relevant to Claim No 13
A	EP,A, 124930 (CROSBY et al) 14 November 1984 (14.11.84) See page 11 lines 8-32 and page 12 lines 1-8	
A	AU,A, 18194/83 (THE UNIVERSITY OF MELBOURNE) 23 February 1984 (23.02.84) See page 3 lines 6-24	
A	AU,A, 46815/85 (THE UNIVERSITY OF MELBOURNE) 13 March 1986 (13.03.86) See page 5 lines 15-29 and page 6 lines 1-24	
A	US,A, 4593696 (HOCHMAIR et al) 10 June 1986 (10.06.86) See abstract	
A	US,A, 4611598 (HORTMANN et al) 16 September 1986 (16.09.86) See abstract	
(continued)		
<p>* Special categories of cited documents: 10</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"&" document member of the same patent family</p>		
IV. CERTIFICATION		
Date of the Actual Completion of the International Search 15 November 1990 (15.11.90)	Date of Mailing of this International Search Report 21 November 1990	
International Searching Authority Australian Patent Office	Signature of Authorized Officer  R. MURRAY	

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

A	US,A, 4063048 (KISSIAH) 13 December 1977 (13.12.77) See abstract
A	US,A, 4592359 (GALBRAITH) 3 June 1986 (03.06.86) See abstract

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE 1

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim numbers ..., because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim numbers , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim numbers ..., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4 (a):

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING 2

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.
2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

- ☐ The additional search fees were accompanied by applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON
INTERNATIONAL APPLICATION NO. PCT/AU 90/00407

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Members		
EP 124930	AU 26683/84	AU 73227/87	AU 73251/87	
	CA 1212501	EP 241101	EP 247649	
	JP 61001200	US 4532930		
AU 46815/85		US 4947844		
US 4592359	AU 55350/86	AU 54985/90	CN 86103049	
	EP 198618	JP 61293475		
US 4611598	DE 3420244	EP 163137		
US 4593696	AU 51720/85	BR 8600156	CA 1247227	
	CN 86100219	EP 190836	JP 61170457	

END OF ANNEX

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.